



Grout, P. A., Mitraile, S. Y., & Sonderegger, S. M. I. A. (2015). The Costs and Benefits of Coordinating with a Different Group. *Journal of Economic Theory*, 160, 517-535.
<https://doi.org/10.1016/j.jet.2015.09.006>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.jet.2015.09.006](https://doi.org/10.1016/j.jet.2015.09.006)

[Link to publication record in Explore Bristol Research](#)
PDF-document

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

The Costs and Benefits of Coordinating with a Different Group

Paul A. Grout^a

University of Bristol and CMPO

Sébastien Mitraille^b

Université de Toulouse, Toulouse Business School

Silvia Sonderegger^c

University of Nottingham and CeDEx

Abstract

We consider a setup where agents care about i) taking actions that are close to their preferences, and ii) coordinating with others. The preferences of agents in the same group are drawn from the same distribution. Each individual is exogenously matched with other agents randomly selected from the population. Starting from an environment where everyone belongs to the same group, we show that introducing agents from a different group (whose preferences are uncorrelated with those of each of the incumbents) generates costs but may also (surprisingly) generate benefits in the form of enhanced coordination.

JEL Codes: C72, D82, Z1. **Keywords:** diversity, coordination, social interactions, value of information, complementarities.

^aCenter for Market and Public Organisation, The Priory Road Complex, Priory Road, Bristol BS8 1TU, United Kingdom.

^bToulouse Business School, 20 BD Lascrosses, 31000 Toulouse, France.

^c**Corresponding author.** E-mail: silvia.sonderegger@nottingham.ac.uk. Address: School of Economics, University of Nottingham, Nottingham NG7 2NR, United Kingdom.

1 Introduction

We consider a setup where individual decision making involves a trade-off between adaptation and coordination.¹ Agents care about i) performing actions that are close to their preferences, and ii) coordinating with others. In our running example, we consider managers who decide how to organize and run production within their divisions. This includes deciding how to assign tasks to workers, which technologies to use, how much emphasis to put on meeting deadlines, etc.² Each manager has preferences over the way in which production should be organized in his division. Different divisions must interact in order to complete a project/task. As in Dessein and Santos (2006) and Alonso *et al.* (2008), we assume that coordination facilitates production and, as a result, managers are concerned with coordinating with those divisions with whom they are matched. Although we focus on production, it is clear that the setup could also be used in other contexts. One is social exchange, as in Kuran and Sandholm (2008); this is affected by personal preferences on how the exchange should occur (dress code, etiquette, etc.) and also by the need to coordinate with others. Another is political activism, as in Dewan and Myatt (2008).³ A crucial feature of our setup is that managers suffer from an information problem: their privately observed preferences are made of two components, a group-specific component and an idiosyncratic component, but they are unable to distinguish between the two. We may think of managers in the same group as belonging to the same school of thought (e.g., Quantitative Approach versus Human Relations Approach) or as having similar expertise, although their precise preferences also have an individual component. Starting from a homogeneous environment (where all managers belong to the same group), we consider the effect of introducing in the organization managers belonging to a different group, whose preferences are uncorrelated with those of each of the incumbents.

The model is introduced in section 2. In section 3, we consider the benchmark case

¹This trade-off has long been recognized as important, not only within economics but also in other disciplines – see e.g. March (1991) for an early contribution in the management literature.

²In this simple example workers are left entirely unmodelled. The way to think about them is as automata that simply follow their division manager’s instructions.

³A further illustration is given by a parent who must instil moral values in his child. He faces a trade-off between the desire to transmit values that reflect his personal preferences and the desire to conform to society at large. E.g., a prudish parent may have a personal preference for instilling a strict moral code in his child, but he may also have to make concessions if other children are raised more liberally, or else his child risks being ostracized by other children. Adriani and Sonderegger (2009) present a model of intergenerational transmission where the coordination concerns of parents arise endogenously.

where all managers in the organization belong to the same group. The key observation here is that manager preferences may differ, even within the same group, because of the presence of idiosyncratic shocks. This translates into within-group heterogeneity of behavior.⁴ Since the environment is characterized by coordination externalities, we show that, from a central planner’s perspective, the equilibrium exhibits too little coordination. This arises because managers put too much weight on their personal preferences when selecting how to organize production in their division.

In section 4, we characterize what happens when some of the firm’s divisions are assigned to managers belonging to another group. The presence of managers of another group has two effects. First, incumbent managers move their actions away from the mean preferences of their own group, and towards their best estimate of the mean preferences of the other group, thus adopting part of the other group’s behavior.⁵ Second, and most importantly, incumbent managers put less weight on their personal preferences when selecting how to organize their divisions. This is because their own preferences are now a less reliable indicator of the preferences of other managers (since these now include some managers belonging to another group). This partially corrects the inefficiencies arising from coordination externalities. At the outset, it may appear that introducing managers from a different group should generally decrease coordination, since their preferences are drawn from a different distribution. We show that this intuition can be misguided once the managers’ reaction to a different environment is taken into account. The surprising implication is that adding managers whose preferences are uncorrelated with those of the incumbents may actually improve coordination.

In section 5 we look at the induced preferences over the group-composition of managers in the organization. We show that the presence of managers from another group always imposes a cost, since incumbents now have to coordinate with managers whose preferences are less likely to be similar to their own. However, as we have argued, it may also generate a benefit in the form of greater coordination. We characterize the sufficient conditions for the net outcome of these opposing forces to induce managers to favor a diverse organization over

⁴The notion that within-group heterogeneity may be pervasive and substantial has been well documented, see e.g. Inglehart (1997) and Hofstede (2001). See also Bednar *et al.* (2010) for a theoretical model of within-group heterogeneity in which individuals care about coordination and consistency.

⁵This shares similarities with Kuran and Sandholm (2008). A tangible example of this phenomenon within the context of social exchange is the appearance of hybrid cuisines and culinary habits in ethnically mixed communities (think for instance of “halal pizza”).

a fully homogeneous one (i.e., where all managers belong to the same group). We also show that, in many cases, preferences exhibit “decreasing returns to scale,” in that incumbents may welcome the presence of managers belonging to another group so long as their share in the organization is sufficiently small, but not otherwise.⁶

Finally, in section 6 we augment the utility function to allow for the possibility that managers may enjoy a direct positive externality from operating in a well-coordinated organization (as in Bolton *et al.* 2010, 2013), and show that this may strengthen the case for introducing diversity in the organization. Section 7 concludes.

In addition to the works cited above, our paper contributes to several literatures. First, it is related to the literature on the composition of teams, such as Prat (2002), Hong and Page (2001) and Lazear (1999). With respect to these works, we present a novel rationale for why diversity may be useful. Second, the paper also contributes to the literature on discrimination. We extend existing work such as Becker (1957), Arrow (1973) and Alesina and La Ferrara (2000, 2002, 2005), in that individual preferences over the composition of their surroundings are derived endogenously. Finally, our approach borrows the tools of the value of information literature –such as for instance Morris and Shin (2002, 2005), Angeletos and Pavan (2004, 2007(a) and (b)), Myatt and Wallace (2012) and Colombo *et al.* (2014)– and the related literature on leadership –Dewan and Myatt (2008), Bolton *et al.* (2010, 2013)– and applies them to a novel set of questions. Our preference-based group structure contrasts with the information-based group structure in Cornand and Heinemann (2008). In a beauty-contest setup where individuals receive a private and a public signal on the fundamental, they show that introducing agents who don’t have access to the public signal (e.g., because its publicity has been curtailed) reduces coordination and may thus increase welfare. By contrast, we show that introducing agents whose preferences are uncorrelated with those of the incumbents may actually improve coordination. Interestingly, the coordination-enhancing effect of diversity does not rely on the information about the other group being public. We show that this effect emerges even when each individual observes an independently drawn private signal about the mean preferences of the other group.⁷

⁶This is reminiscent of observed attitudes towards newcomers, such as immigrants. Quillian (1995), controlling for individuals factors, shows the existence of a positive correlation between population size of the racial minority and the degree of racial prejudices expressed by natives of the country.

⁷See footnotes 12 and 17 below for further elaborations of this point.

2 Model

Background and utility. An organization of measure 1 contains a continuum of divisions each headed by a manager. Each manager, and hence each division, is represented by a real coordinate i on the unit interval $[0, 1]$. Managers decide how to organize production within their division. Each division is randomly matched with $n \in \mathbb{N}_1$ other divisions in an interacting unit that must collaborate on a task.⁸ The utility of a generic manager i who selects an action $a_i \in \mathbb{R}$ is $-U(a_i - \bar{a}_i, a_i - \alpha_i)$, where

$$U(a_i - \bar{a}_i, a_i - \alpha_i) = \theta (a_i - \bar{a}_i)^2 + (1 - \theta) (a_i - \alpha_i)^2 \quad (1)$$

for some $\theta \in (0, 1)$; $\bar{a}_i = \frac{1}{n} \sum_{j=1}^n a_j$ is the average action of the n managers heading the divisions with whom i 's division is matched and $\alpha_i \in \mathbb{R}$ is a taste parameter that reflects i 's preferences.⁹ This captures the idea that the costs incurred by the manager to complete the task depend on 1) how closely he can adhere to his own preferences/expertise when organizing production in his division and, and 2) how coordinated his division is with the other divisions with whom it collaborates on the task. Throughout the analysis we will confine attention to finite values of n , unless otherwise stated. A few points are worth discussing.

First, we do not allow managers to adjust their actions to suit the realized composition of their interacting unit. Intuitively, managers cannot constantly change how their divisions operate, but rather they must instruct their workers on a general line of conduct before they know the identity of the divisions with whom they will collaborate on a task. This feature is plausible also in other applications. For instance, a parent instilling values in his child does so before the child grows up and interacts with others. Moreover, although adjusting one's behavior to different encounters may potentially be useful, it may also be prohibitively costly – this point is also raised by Kuran and Sandholm (2008, pp.203-204).

Second, we assume that managers cannot communicate with each other before they select their actions. This assumption is standard in the value of information literature – see e.g.

⁸Interacting units are composed of n equally spaced coordinates on $[0, 1]$. The reader may refer to the supplementary material. Note that the composition of each interacting unit is exogenously given. This is well suited to environments (such as for instance the workplace, as in our running example) where agents cannot actively choose their counterparties. An obvious extension would consider endogenous assortment. We leave the analysis of this possibility to future research.

⁹In what follows, the term “preferences” will be utilized to indicate α_i , unless otherwise stated. Note that our model could also accommodate alternative interpretations of α_i . For instance, α_i could capture individual i 's identity or self-image, as in Akerlof and Kranton (2000).

Angeletos and Pavan (2007a) – and in the optimal organizational design literature – see e.g., Prat (2002) – as it allows to abstract from strategic information transmission and to concentrate on issues purely connected with coordination. In many instances it is also a close approximation of reality.

Last, in the model we consider the manager’s production cost depends only the distance between his action and his preferences and on the distance between his action and the weighted average of the actions of others, *not* on the action itself. This allows us to concentrate on effects arising from coordination, and to abstract from situations where one group has, on average, inherently “superior” preferences.

Groups. Managers are divided in two groups, A and B . The share of each group in the population is common knowledge: the proportion of managers of group $t = A, B$ is equal to λ^t , with $\lambda^A + \lambda^B = 1$. We assume that manager preferences are given by the sum of two components: a group-specific component, equal to $\mu^t + e^t$, and an idiosyncratic component, equal to ε_i . The group-specific component $\mu^t + e^t$ corresponds to the sum of a mean preference μ^t and a random element e^t , common to all managers of group t . We can think of e^t as capturing the effect of specific circumstances on the group-specific component of preferences, while μ^t corresponds to mean preferences when specific circumstances are averaged out.

To sum up, therefore, the taste parameter α_i^t of a manager i belonging to group t is equal to

$$\alpha_i^t = \mu^t + e^t + \varepsilon_i \quad (2)$$

where for any group $t \in \{A, B\}$, μ^t represents the mean preferences of group t , e^t drawn as $N(0, 1)$ represents the group-specific shock to preferences, with $e^A \perp e^B$, and ε_i , the idiosyncratic shock to preferences, is drawn as $N(0, \sigma^t)$ for a positive constant σ^t , with $\varepsilon_i \perp \varepsilon_j$ for $i \neq j$, where $\varepsilon_i \perp e^t$ for any i and t .¹⁰ For ease of exposition the superscript t will be dropped from α_i^t whenever this does not cause any ambiguity.¹¹

¹⁰Since our analysis focuses on the implications of information imperfections (rather than intrinsic asymmetries between the groups) for the desirability of diversity, our case of interest is when the two groups are symmetric, i.e. $\sigma^A = \sigma^B$. However, we will maintain the notation σ^A and σ^B as much as possible, in order to separate the effect of a change in the variance of the idiosyncratic shock in one’s own group (respectively, in the other group) on a manager’s actions and preferences for diversity.

¹¹Moreover, note that the distribution of ε_i is strictly speaking a function of the group t of agent i through its variance σ^t . We omit the superscript t for ease of exposition.

Information. Manager preferences are private information. Each manager observes his own α_i^t , but he is unable to discriminate between e^t and ε_i , the group-specific and the idiosyncratic shocks affecting his preferences. Formally, consider manager i belonging to group $t_i = t$, and a randomly drawn manager $j \neq i$ who also belongs to group t . From i 's perspective, $E(\alpha_j^t | t_i, \alpha_i, x_i) = \mu^t + E(e^t | t_i, \alpha_i, x_i) + E(\varepsilon_j | t_i, \alpha_i, x_i) = \mu^t + E(e^t | t_i, \alpha_i, x_i)$ (since $E(\varepsilon_j | t_i, \alpha_i, x_i) = 0$). The prediction of the group-specific shock to the preferences of group t managers, e^t , is given by the linear regression of e^t against all the relevant information available to manager i (see Morris and Shin 2002 or Angeletos and Pavan 2007a). Hence, $E(e^t | t_i, \alpha_i, x_i) = (\alpha_i - \mu^t) / (\sigma^t + 1)$ and thus, $E(\alpha_j^t | t_i, \alpha_i, x_i) = (\mu^t \sigma^t + \alpha_i) / (\sigma^t + 1)$.

We assume that all group $t \in \{A, B\}$ managers know the value of μ^t . However, their beliefs over $\mu^{t'}$ are dispersed: each manager i only receives a signal x_i^t over it. Beliefs are independently and identically distributed as a Normal distribution with unknown mean $\mu^{t'}$ and variance X , $x_i^t \sim N(\mu^{t'}, X)$ for some $X > 0$, with $x_i^t \perp x_j^t$ for $i \neq j$.¹² This implies that, for each manager i of group t , the best estimate of $\mu^{t'}$ is given by his private signal x_i^t . In what follows, we will at time use the notation $\xi_i \equiv x_i^t - \mu^{t'}$, so that $x_i^t = \mu^{t'} + \xi_i$. For ease of exposition the superscript t will be dropped from x_i^t whenever this does not cause any ambiguity. Finally, we assume that $\alpha_i^t \perp x_i^t$.

Timing. The timing of the game is as follows.

$t = 0$ The group-specific shocks e^A , e^B and the individual shocks ε_i and ξ_i are realized.¹³

$t = 1$ Each division manager observes α_i and x_i and selects how to organize production in his division.

$t = 2$ Divisions are randomly matched in interacting units.¹⁴

$t = 3$ Payoffs are realized.

¹²As will become clear below, the restriction $X > 0$ is not instrumental to our results, all of which (except one, i.e., proposition 6) require that X should not be too large. However, allowing for $X > 0$ (i) clarifies that our results are not purely confined to $X = 0$ (i.e., the information about the other group is public) and (ii) emphasizes that the result of proposition 6 holds independently of X , which is an interesting observation in its own right. We thank George-Marios Angeletos for pointing this out.

¹³We are assuming that the same group-specific shocks affect the whole manager population. Alternatively, we could have modeled these shocks as being independently drawn within each interacting unit. This would have not affected our results.

¹⁴In what follows, for brevity, we will at times refer to the “managers with whom i is matched” to mean the “managers who head the divisions that belong to the same interacting unit as i 's division.”

3 Homogeneous environment

We start off by studying the benchmark case of an homogeneous environment, in which all managers belong to the same group. In what follows, we accordingly drop the superscript t . The first step of our analysis consists in identifying the (constrained) efficient decision rule, which can then be compared to what actually happens in equilibrium.

The social planner's solution. Consider the problem of a social planner, who can decide which decision rule must be used by all managers in order to maximize their joint payoff (the Decentralized Information Optimum, as in Hellwig 2005 or Angeletos and Pavan 2007a).

The social planner solves $\min_k E[U(a_i - \alpha_i, a_i - \bar{a}_i)]$ subject to

$$a_i = k\alpha_i + (1 - k)\mu \quad \forall i. \quad (3)$$

Lemma 1. (*Social planner's solution.*) *The solution to the social planner's problem is*

$$a_i = k^*\alpha_i + (1 - k^*)\mu \quad (4)$$

where $k^* \equiv \frac{(1-\theta)(\sigma+1)}{\sigma+1-\theta(1-\frac{\sigma}{n})}$.

Proof: See appendix. ■

Equilibrium. The equilibrium action solves $\min_{a_i} E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid \alpha_i]$.

Lemma 2. (*Equilibrium in homogeneous environment.*) *In the unique equilibrium, the action of manager i is*

$$a_i^t = k^e\alpha_i + (1 - k^e)\mu \quad (5)$$

where $k^e \equiv \frac{(1-\theta)(\sigma+1)}{\sigma+1-\theta}$.

Proof: The proof is an adaptation of that of lemma 3 to the case $\lambda^{t'} = 0$. ■

Corollary 1. *In equilibrium managers put too much weight on their personal preferences compared with the social optimum: $k^e > k^*$ for all finite n .*

Intuitively, the environment is characterized by coordination externalities: when choosing their actions, managers fail to take into account that they will affect the coordination costs of other managers. From a central planner's perspective, this generates too little coordination, since managers put too much weight on their personal preferences when selecting how to organize their divisions.

Inspection of k^* and k^e reveals that equilibrium behavior becomes optimal only in the limit case $n \rightarrow \infty$: if interaction units are modelled as containing a continuum of divisions, then each individual division has zero mass, and, consequently, the coordination externality described above disappears.

4 Equilibrium with heterogeneity

We now move away from homogeneous environments and provide a full characterization of the equilibrium of the game. Letting t_i denote i 's group, the equilibrium action solves $\min_{a_i} E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t_i, \alpha_i, x_i]$. This gives

$$a_i = \theta \bar{\bar{a}}_i + (1 - \theta) \alpha_i \quad (6)$$

where $\bar{\bar{a}}_i \equiv E(\bar{a}_i \mid t_i, \alpha_i, x_i)$ is manager i 's expectation of the average action of the managers with whom i is matched.

Lemma 3. (*Description of equilibrium.*) *In the unique equilibrium, the action of manager i of group $t \in \{A, B\}$ is given by*

$$a_i^t = k^t \alpha_i + \theta \lambda^{t'} x_i + (1 - k^t - \theta \lambda^{t'}) \mu^t \quad (7)$$

where $k^t \equiv \frac{(1-\theta)(\sigma^t+1)}{\sigma^t+1-\theta(1-\lambda^{t'})}$.

Proof: See appendix. ■

Lemma 3 describes the equilibrium strategy followed by managers, as a function of the composition of the organization. Each manager selects an action that is a weighted average of his preferences, the mean preferences of managers of his same group (when specific circumstances are averaged out), and his forecast of the mean preferences of managers of the other group.

Expression (7) clarifies the effect that the presence of managers of group t' has on the behavior of a manager of group t . First, it induces him to move away from μ^t and towards x_i – namely, his best estimate of $\mu^{t'}$. Second, the presence of group t' also affects k^t , namely the weight managers of group t put on their personal preferences when deciding how to organize production in their division. The key feature here is that, given $\theta < 1$, k^t is strictly *decreasing* in $\lambda^{t'}$, the proportion of managers of group t' . Hence, the presence of managers of a different group counteracts the inefficiency identified in section 3, which takes the form of managers putting too much weight on their personal preferences when selecting their actions.

To see why k^t decreases in $\lambda^{t'}$ note that a manager's expectation of the average action in his interacting unit depends on his estimate of average preferences. It is here that the group-composition of management within the organization makes a difference. This is because the estimate that manager i of group t can make of the preferences of a group t' manager can only

be based on x_i , his noisy private signal on $\mu^{t'}$. By contrast, when predicting the preferences of another manager of his own group t , i is better equipped, since he knows μ^t and he also observes his preference parameter α_i^t . This latter piece of information matters because, from (2), a manager's personal preferences are correlated with the preferences of others of the same group. Hence, in equilibrium, the manager's behavior depends on α_i for two reasons: (i) the desire to accommodate his personal preferences, and (ii) the concern for coordination with other managers of the same group. Clearly enough, this latter motive depends on how widespread the manager's group is within the firm. If it is very widespread, then his interacting unit is likely to contain many other managers from that group, and, consequently, the manager's concern for coordination with others from his same group is strengthened (and vice-versa). As a result, the weight put on α_i by the manager's equilibrium action is a decreasing function of $\lambda^{t'}$, the share of managers of the other group. This observation is key to understanding our results.

5 Induced preferences over the composition of management in the organization

We now look at the utility a manager can expect to obtain when working in a given organization. We consider the ex-ante expectation $E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t_i, x_i]$, computed *behind a veil of ignorance*, namely, before α_i is realized. The idea is that, at the ex-ante stage, a manager does not know what his exact preferences will be, since these depend on circumstances that have yet to realize.¹⁵

Lemma 4. *(A useful decomposition.) A manager's ex-ante expected (indirect) utility can be written as*

$$-E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t_i, x_i] = -\varphi_0 E[(\bar{a}_i - \alpha_i)^2 \mid t_i, x_i] - \varphi_1 E[(\bar{a}_i - \bar{a}_i)^2 \mid t_i, x_i] \quad (8)$$

where $\varphi_0 \equiv 0.5\theta(1 - \theta)$ and $\varphi_1 \equiv 0.5\theta$.

Proof: See appendix. ■

Lemma 4 establishes that a manager's expected utility can be expressed as the weighted sum of two components, which have intuitive interpretations. Consider the first element of (8), namely $E[(\bar{a}_i - \alpha_i)^2]$ (we omit the conditioning on t_i, x_i for brevity). Keeping everything

¹⁵Note that we let the ex-ante expectation be conditional on x_i , which allows us to characterize expected utility as a function of a manager's beliefs over the other group's preferences. However, it is straightforward to see that this plays no role for our results.

else equal, managers dislike environments where they expect that, on average, production will be organized in a way that differs from their personal preferences. Consider now the second element in (8), namely $E[(\bar{a}_i - \bar{a}_i)^2]$. This is the variance of \bar{a}_i around its conditional mean, $\bar{\bar{a}}_i$. A small value of $E[(\bar{a}_i - \bar{a}_i)^2]$ implies that i 's expectation is a good predictor of the actual average action in his interacting unit. In other words, manager i faces low *strategic uncertainty*. Low strategic uncertainty raises expected utility, since it makes it easier to coordinate with other divisions.

Proposition 1. (*Cost of Diversity.*) For a manager i of group t , $E[(\bar{a}_i - \alpha_i)^2 \mid t, x_i]$ is always minimized at $\lambda^{t'} = 0$.

Proof: See appendix. ■

Intuitively, this follows because managers of the same type are more likely to have similar preferences. Things can however change once we introduce the second component of (8), namely strategic uncertainty. By lowering strategic uncertainty, the presence of managers of a different group may well improve individual welfare. We now identify the sufficient conditions for this to be the case.

Proposition 2. (*Individual preferences.*) Consider a manager i of group t and denote $|x_i - \mu^t|$ as z_i . There exist $X' > 0$, $z' > 0$, $\Delta' > 0$ and $\theta' \in [0, 1]$ (all independent of n) such that, when $X < X'$, $z_i < z'$, $\sigma^{t'} - \sigma^t < \Delta'$ and $\theta \in (\theta', 1)$, $E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t, x_i]$ is minimized at $\lambda^{t'} > 0$.

Proof: See appendix. ■

Proposition 2 shows that diversity may actually end up lowering strategic uncertainty, and that this effect may be sufficiently strong to ensure that a manager's expected utility is maximized in a mixed rather than a homogeneous organization. The sufficient conditions for this to occur are that z_i , X and $\sigma^{t'} - \sigma^t$ should not be too large, and θ should not be too small. Intuitively, a large z_i suggests to manager i that the mean preferences of the two groups are very different (recall that $z_i \equiv |x_i - \mu^t|$). Keeping everything else equal, this increases the dispersion of \bar{a}_i around $\bar{\bar{a}}_i$, its expected value.¹⁶ Consider now $\sigma^{t'}$. Keeping everything else equal, a larger $\sigma^{t'}$ increases the dispersion of group t' 's actions, making the behavior of managers belonging to this group harder to predict. For a t individual, introducing group

¹⁶ Consider for instance the simple case where $n = 1$. In that case, $\bar{\bar{a}}_i = \lambda^{t'} E(a_j^{t'} \mid \alpha_i, x_i) + (1 - \lambda^{t'}) E(a_j^t \mid \alpha_i, x_i)$, but \bar{a}_i is equal to *either* $a_j^{t'}$ (with probability $\lambda^{t'}$) *or* a_j^t (with probability λ^t). If i believes that the two groups have very different mean preferences, this clearly increases the distance between $\bar{\bar{a}}_i$ and $a_j^{t'}$ or a_j^t .

t' people in the organization may lower strategic uncertainty only if $\sigma^{t'}$ is not too large compared to σ^t . Similarly, a large X also makes managers belonging to the other group harder to predict, and thus raises strategic uncertainty. Moreover, a large X makes the managers's predictions on group t' more dispersed within group t ; keeping everything else equal, this makes other managers of the same group harder to predict.¹⁷ Finally, the last condition of proposition 2 is that θ should not be too small. This is because a small θ lowers the concern for coordination and therefore makes managers less responsive to changes in organizational composition.

Although a full characterization of the “ideal” $\lambda^{t'}$ cannot be achieved due to intractability of the analytical expressions, a numerical illustration may be useful. Throughout the paper, we will focus on the following example: $\sigma^A = \sigma^B = 1$, $\theta = 0.75$, $n = 1$ and $z_i = 0.25$. The numerical analysis makes clear that the beneficial effect of diversity may be quite sizeable. In our example, for instance, when $X = 0$ a group t manager's utility maximizing share of group t' is 9%.¹⁸

Finally, note that Proposition 2 holds for all finite n . Things would however change if we were to consider the limit $n \rightarrow \infty$.

Proposition 3. *In the limiting case where $n \rightarrow \infty$, $E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t, x_i]$ is minimized at $\lambda^{t'} = 0$.*

Proof: See appendix. ■

Proposition 3 shows that if equilibrium behavior in the homogeneous environment coincides with the social planner's solution (which is the case if and only if $n \rightarrow \infty$), then introducing managers of another group is *never* desirable for the incumbents. The desirability of diversity is thus contingent on the existence of inefficiencies in the homogenous environment.

The last proposition in this section argues that attitudes towards managers of a different group may eventually become less welcoming as their share increases.

Proposition 4. *(Decreasing returns.) Consider a manager i of group t . There exist $z' > 0$ and $\theta' \in [0, 1)$ such that, when $z_i < z'$ and $\theta \in (\theta', 1)$, $E[(\bar{a}_i - \bar{a}_i)^2 \mid t, x_i]$ is convex in $\lambda^{t'}$.*

¹⁷Note that, by modelling the signals x_i^t and x_j^t as i.i.d., we are working against what we are set to prove. Consider for instance an alternative scenario, where all managers within the same group $t = \{A, B\}$ receive the *same* signal x^t about $\mu^{t'}$. For a manager of group t , this would make the reaction of others of his same group to the presence of managers of group t' easier to predict, and would thus increase the desirability of diversity. Indeed, it can be easily shown that, if all managers within the same group were to receive the same signal x^t , and assuming $X > 0$, then for any $\lambda^{t'} > 0$ the value of $E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t, x_i]$ would be *smaller* than in the case we analyze. We thank an anonymous referee for pointing this out.

¹⁸We let the reader refer to the next section (Figures 2 and 4) for a more general characterization.

Proof: See appendix. ■

Intuitively, if a manager's group represents a small fraction of the manager population, then he will put little weight on his own preferences when selecting his actions – since these are poor predictors of the actions of other managers. Hence, if a small share of managers of, say, group t' are introduced in an organization composed predominantly of group t , they will put little weight on their personal preferences, and instead put a lot of weight on (their estimate of) the mean preferences of group t . From the perspective of group t , these managers are then easy to predict and to coordinate with. However, as their share increases, managers of group t' start to put more weight on their personal preferences when organizing production in their divisions. From the point of view of the other group, this makes them harder to predict and thus decreases their appeal. Note that the sufficient conditions for decreasing returns are consistent with those in Proposition 2.

6 Positive externality from a well-coordinated organization

We now allow for the possibility that managers may enjoy a direct positive externality from working in a well-coordinated organization, as in Bolton *et al.* (2010, 2013). To this purpose, we let their payoff depend on a third element, namely heterogeneity of actions. We consider two variants.

First variant. Suppose that $U(\cdot)$ is given by

$$U(a_i - \bar{a}_i, a_i - \alpha_i, \text{Var}(a_j)) = \theta (a_i - \bar{a}_i)^2 + (1 - \theta) (a_i - \alpha_i)^2 + \zeta \text{Var}(a_j) \quad (9)$$

where $\text{Var}(a_j) \equiv E[(a_j - \bar{a})^2 \mid e^A, e^B]$ is the realized heterogeneity of behavior within the whole organization, $\bar{a} \equiv E(a_j \mid e^A, e^B)$ represents overall average action, and $\zeta > 0$. Note that, since each division is of measure zero within the whole organization, the inclusion of $\text{Var}(a_j)$ leaves the equilibrium action (and, thus, the equilibrium) unchanged from that identified in lemma 3. However, a manager's preferences over organizational composition are now different. To calculate $\text{Var}(a_j)$, we consider a randomly drawn manager, who may belong either to group A or to group B , and we compute the variance of his action conditional on group shocks.¹⁹ We then derive the ex-ante expectation $E[\text{Var}(a_j) \mid t, x_i]$,

¹⁹ Although common shocks are a source of variation from an ex ante viewpoint, from an ex post perspective they are held fixed, and hence they do not affect the *realized* heterogeneity of actions. The fact that we concentrate on conditional variance is however not crucial for the result.

computed behind a veil of ignorance (i.e., before α_i is realized).²⁰

Proposition 5. (*Diversity and overall heterogeneity of actions.*) Consider a manager i of group t . There exist $X' > 0$, $z' > 0$, $\theta' \in [0, 1)$ and $\Delta' > 0$ such that, when $X < X'$, $z_i < z'$, $\sigma^{t'} - \sigma^t < \Delta'$ and $\theta \in (\theta', 1)$ then $E[\text{Var}(a_j) \mid t, x_i]$ is minimized when $\lambda^{t'} > 0$.

Proof: See appendix. ■

Proposition 5 shows that adding to a group that contains within-group heterogeneity a second group whose preferences are uncorrelated with the preferences of each of the incumbents may reduce the overall heterogeneity of actions. When this occurs, the presence of a direct positive externality from working in a well-coordinated organization increases the desirability of diversity.²¹

Figure ?? depicts $E[\text{Var}(a_j) \mid t, x_i]$ as a function of $\lambda^{t'}$ in our numerical example, for different values of X . Figure ?? depicts the utility-maximizing share $\lambda^{t'}$ from the perspective of a manager of group t , for different values of ζ (including $\zeta = 0$, namely, the case considered in the previous section) and X .²²

Second variant. We now explore a second variant. Suppose that $U(\cdot)$ is given by

$$U(a_i - \bar{a}_i, a_i - \alpha_i, \text{Var}(a_j^t)) = \theta (a_i - \bar{a}_i)^2 + (1 - \theta) (a_i - \alpha_i)^2 + \zeta \text{Var}(a_j^t) \quad (10)$$

where $\text{Var}(a_j^t) \equiv E[(a_j^t - \bar{a}^t)^2 \mid e^A, e^B]$ is the realized heterogeneity of actions *within group* t , $\bar{a}^t \equiv E(a_j^t \mid e^A, e^B)$ represents average action in group t , and $\zeta > 0$. This variant is not particularly well-suited to our running example of production within a firm, but may apply quite well to social exchanges, where it is easy to see that a given cultural group may enjoy direct benefits from being well coordinated. In what follows, we accordingly drop the “manager/division” terminology and instead refer to individuals interacting with one another. Since each individual is of measure zero within his whole group, the equilibrium

²⁰The reason why we take ex-ante expectations is that $\text{Var}(a_j)$, the realized heterogeneity of actions, depends on the precise realizations of e^A and e^B .

²¹Note that the conditions laid out in proposition 5 are qualitatively similar to those identified in proposition 2. This is because in both cases the coordination-enhancing effect of diversity is the driving force for the result.

²²Generally, the precise utility-maximizing composition of management within the organization will differ between the two groups. However, this does not preclude that both groups may jointly favor diversity over homogeneity. Consider for instance the numerical example discussed above, and suppose that $\zeta = 1$ and $X < 0.15$. Then it is easy to verify that managers from both groups prefer an environment in which the share of each group is 50% to a homogeneous environment (in which all managers belong to their same group). Straightforward calculations show that this is also true in the second variant (below), except that the restriction on X in that case becomes $X < 0.16$.

action (and, thus, the equilibrium) is unaffected by the presence of $Var(a_j^t)$.²³

Proposition 6. (*Diversity and within-group heterogeneity of actions.*) *Introducing a few individuals of the other group lowers within-group heterogeneity of actions at the margin: evaluated at $\lambda^{t'} \rightarrow 0$, $Var(a_j^t)$ is strictly decreasing in $\lambda^{t'}$ for all values of X . This also holds in the limit case where $n \rightarrow \infty$.*

Proof: Substituting for a_j^t from (7), and since $\bar{a}^t = k^t e^t + \theta \lambda^{t'} \mu^{t'} + (1 - \theta \lambda^{t'}) \mu^t$, we have $a_j^t - \bar{a}^t = k^t \varepsilon_j + \theta \lambda^{t'} \xi_j$, and, hence, $E[(a_j^t - \bar{a}^t)^2 | e^A, e^B] = (k^t)^2 \sigma^t + (\theta \lambda^{t'})^2 X$. In the limit $\lambda^{t'} \rightarrow 0$ the derivative of $Var(a_j^t)$ w.r.t. $\lambda^{t'}$ is equal to $-2\theta \sigma^t (\sigma^t + 1)^2 (\theta - 1)^2 / (1 - \theta + \sigma^t)^3 < 0$.

■

The impact of a larger $\lambda^{t'}$ on the heterogeneity of the actions of group t is *a priori* ambiguous. On the one hand, it reduces k^t . By inspection of (7), it is clear that a smaller k^t implies that actions are less affected by the realizations of ε_i , namely the idiosyncratic shocks to preferences. This reduces the heterogeneity of actions of group t individuals. On the other hand, t individuals now put more weight on their private signal x_i . Keeping everything else equal, this makes their actions more dispersed. For $\lambda^{t'}$ sufficiently small, however, the second effect always becomes second-order and, thus, the first effect dominates. Intuitively, this follows because the weight that individual actions put on x_i is directly proportional to the share of the other group in the population, and thus becomes negligible when this share is sufficiently small. Hence, the result holds independently of the value of X , namely the dispersion of beliefs x_i on the other group's preferences. An implication of Proposition 6 is that,

Corollary 2. *Consider an individual i of group t . There exist $\zeta' > 0$ such that, when $\zeta > \zeta'$, $E[U(.) | t, x_i]$ is minimized at $\lambda^{t'} > 0$.*

Figure ?? depicts $Var(a_j^t)$ as a function of $\lambda^{t'}$ in our numerical example, for different values of X . Figure ?? depicts the utility-maximizing share $\lambda^{t'}$ from the perspective of an individual of group t , for different values of ζ and X .

7 Concluding remarks

Our analysis underscores the two-way relationship between individual behavior and environment: on the one hand, each individual shapes the environment he belongs to; on the other

²³Note that, since $E[(a_j^t - \bar{a}^t)^2 | e^A, e^B]$ is independent of e^A and e^B , taking ex-ante expectations (as we did in the first variant) is unnecessary in this case.

hand, the type of environment an individual faces shapes the way he behaves. This latter effect ensures that adding individuals of a different group improves welfare, since it induces the incumbents to rely less on personal preferences when selecting their actions.

There are a number of ways in which the analysis can be advanced. One is endogenous matching. Another is competition among firms. Our analysis suggests that by introducing some diversity in their workforce firms may become more attractive to potential employees. What does the equilibrium look like in that case? We leave these questions to future research.

Acknowledgements

We thank Fabrizio Adriani, George-Marios Angeletos, Steven Durlauf, Miltos Makris, Eric Maskin, Peyton Young and seminar participants at the Universities of Birmingham, Bolzano-Bozen, Bristol, Exeter, Halifax, Keele, Lausanne, Louvain-la-Neuve, Milan, Royal Holloway and Vienna for useful comments and discussions. We are also indebted to the Editor, Christian Hellwig, and two anonymous referees for valuable suggestions. All errors are our own.

References

- [1] Adriani, F., and S. Sonderegger (2009) Why Do Parents Socialize Their Children To Behave Pro-Socially? An Information-Based Theory, *J. Public Econ.* 93(11/12): 1119-1124.
- [2] Akerlof, G. A., and R. Kranton (2000) Economics and Identity, *Q. J. Econ.* 115(3): 715-753.
- [3] Alesina, A., and E. La Ferrara (2000) Participation in Heterogeneous Communities, *Q. J. Econ.* 115(3): 847-904.
- [4] Alesina, A., and E. La Ferrara (2002) Who Trusts Others?, *J. Public Econ.* 85(2): 207-234.
- [5] Alesina, A., and E. La Ferrara (2005) Ethnic Diversity and Economic Performance, *J. Econ. Lit.* 43(3): 762-800.
- [6] Alonso, R., W. Dessein, and N. Matouschek (2008) When does Coordination require Centralization?, *Am. Econ. Rev.* 98(1): 145-179.

- [7] Angeletos, G. M., and A. Pavan (2004) Transparency of Information and Coordination in Economies with Investment Complementarities, *Am. Econ. Rev.* 94(2): 91-98.
- [8] Angeletos, G. M., and A. Pavan (2007a) Efficient Use of Information and Social Value of Information, *Econometrica* 7(4): 1103-1142.
- [9] Angeletos, G. M., and A. Pavan (2007b) Socially Optimal Coordination: Characterization and Policy Implications, *J. Eur. Econ. Assoc.* 5(2-3): 585-593.
- [10] Arrow, K. J. (1973) *The Theory of Discrimination*, in: O. Ashenfelter and A. Rees (Eds.), *Discrimination in Labor Markets*, Princeton University Press, Princeton (NJ), 1973.
- [11] Becker, G. S. (1957) *The Economics of Discrimination*, University of Chicago Press, Chicago (IL), 1957.
- [12] Bednar, J., A. Bramson, A. Jones-Rooy, and S. E. Page (2010) Emergent Cultural Signatures and Persistent Diversity: A Model of Conformity and Consistency, *Ration. Soc.* 22(4): 407-444.
- [13] Bolton, P., M. Brunnermeier, and L. Veldkamp (2010) *Economists' Perspectives on Leadership*, in: N. Nohria and R. Khurana (Eds.), *Handbook of Leadership Theory and Practice*, Harvard Business Press, Boston (MA), 2010.
- [14] Bolton, P., M. Brunnermeier, and L. Veldkamp (2013) Leadership, Coordination and Mission-Driven Management, *Rev. Econ. Stud.* 80(2): 512-537.
- [15] Colombo, L., G. Femminis, and A. Pavan (2014) Information Acquisition and Welfare, *Rev. Econ. Stud.* 81(4): 1187-1483.
- [16] Cornand, C., and F. Heinemann (2008) Optimal Degree of Public Information Dissemination, *Econ. J.* 118(528): 718-742.
- [17] Dessein, W., and T. Santos (2006) Adaptive Organizations, *J. Polit. Econ.* 114(5): 956-995.
- [18] Dewan, T., and D. P. Myatt (2008) The qualities of leadership: Direction, communication, and obfuscation, *Am. Polit. Sci. Rev.* 102(3): 351-368.

- [19] Hellwig, C. (2005) Heterogeneous Information and the Welfare Effects of public information Disclosures, mimeo.
- [20] Hofstede, G. (2001) *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*, Sage Publications, Thousand Oaks (CA), 2001.
- [21] Hong, L., and S.E. Page (2001) Problem Solving by Heterogeneous Agents, *J. Econ. Theory* 97(1): 123-163.
- [22] Inglehart, R. (1997) *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*, Princeton University Press, Princeton (NJ), 1997.
- [23] Kuran, T., and W. H. Sandholm (2008) Cultural Integration and Its Discontents, *Rev. Econ. Stud.* 75(1): 201-228.
- [24] Lazear, E. P. (1999) Culture and Language, *J. Polit. Econ.* 107(6): S95-S126.
- [25] March, J. G. (1991) Exploration and Exploitation in Organizational Learning, *Organ. Sci.* 2: 71-87.
- [26] Morris, S., and H. S. Shin (2002) Social Value of Public Information, *Am. Econ. Rev.* 92(5): 1521-1534.
- [27] Morris, S., and H. S. Shin (2005) *Heterogeneity and Uniqueness in Interaction Games*, in: L. Blume and S. Durlauf (Eds.), *The Economy as an Evolving Complex System III*, Santa Fe Institute Studies in the Sciences of Complexity, Oxford University Press, New York, 2005.
- [28] Myatt, D., and C. Wallace (2012) Endogenous Information Acquisition in Coordination Games, *Rev. Econ. Stud.* 79(1): 340-374.
- [29] Prat, A. (2002) Should a Team Be Homogeneous?, *Eur. Econ. Rev.* 46(7): 1187-1207.
- [30] Quillian, L. (1995) Prejudice as a Response to perceived Group Threat: Population Composition and Anti-Immigrant and Racial prejudice in Europe, *Am. Sociol. Rev.* 60(4): 586-611.

Appendix

Proof of Lemma 1. From (3), $\bar{a}_i = \frac{k}{n} \sum_j \alpha_j + (1-k)\mu$. Substituting for $\alpha_j = \mu + e + \varepsilon_j$ $\forall j$, we obtain $\bar{a}_i = \mu + ke + \frac{k}{n} \sum_j \varepsilon_j$, so that $a_i - \bar{a}_i = k\varepsilon_i - \frac{k}{n} \sum_j \varepsilon_j$. Similarly, $a_i - \alpha_i = (k-1)(e + \varepsilon_i)$. Substituting for these in $U(a_i - \alpha_i, a_i - \bar{a}_i)$ and taking expectations, we obtain $\theta k^2 E[(\varepsilon_i - \frac{1}{n} \sum_j \varepsilon_j)^2] + (1-\theta)(k-1)^2 E[(e + \varepsilon_i)^2]$, i.e., $\theta k^2 \sigma (1 + \frac{1}{n}) + (1-\theta)(k-1)^2 (1 + \sigma)$. This is minimized at $k = k^*$. ■

Proof of Lemma 3. The proof is divided in two parts. First, we characterize the linear equilibrium of the game. Second, we show that the linear equilibrium is also the unique equilibrium of the game. **Linear equilibrium** In a linear equilibrium, the actions of a generic individual j of group $\tau \in \{A, B\}$ can be written as a linear combination of preferences and beliefs, $a_j = k^\tau \alpha_j + \gamma^\tau x_j + \delta^\tau \mu^\tau + s^\tau$. Consider now \bar{a}_i . This is equal to $\frac{1}{n} E(\sum_{j=1}^n a_j \mid t_i, \alpha_i, x_i)$. By linearity of expectation, $\frac{1}{n} E(\sum_{j=1}^n a_j \mid t_i, \alpha_i, x_i) = \frac{1}{n} \sum_{j=1}^n E(a_j \mid t_i, \alpha_i, x_i)$. Since expected actions are identity-independent, it follows that $\bar{a}_i = E(a_j \mid t_i, \alpha_i, x_i)$ for a generic $j \neq i$. Note that

$$E(a_j \mid t_i, \alpha_i, x_i) = \sum_{\tau=A,B} \lambda^\tau [k^\tau E(\alpha_j^\tau \mid t_i, \alpha_i, x_i) + \gamma^\tau E(x_j^\tau \mid t_i, \alpha_i, x_i) + \delta^\tau E(\mu^\tau \mid t_i, \alpha_i, x_i) + s^\tau]. \quad (11)$$

Suppose now that $t_i = t$ and let $t' = \{A, B\} \setminus t$. Substituting for $E(\alpha_j^{t'} \mid t, \alpha_i, x_i) = E(\mu^{t'} \mid t, \alpha_i, x_i) = E(x_j^t \mid t, \alpha_i, x_i) = x_i$, $E(x_j^{t'} \mid t, \alpha_i, x_i) = \mu^t$ and $E(\alpha_j^t \mid t, \alpha_i, x_i) = (\sigma^t \mu^t + \alpha_i)/(\sigma^t + 1)$ in (11) gives

$$E(a_j \mid t, \alpha_i, x_i) = \lambda^{t'} (k^{t'} x_i + \gamma^{t'} \mu^t + \delta^{t'} x_i + s^{t'}) + (1 - \lambda^{t'}) (k^t \frac{\sigma^t \mu^t + \alpha_i}{\sigma^t + 1} + \gamma^t x_i + \delta^t \mu^t + s^t). \quad (12)$$

Substituting for (12) in the first-order condition (6) we obtain

$$\begin{aligned} k^t &= \theta(1 - \lambda^{t'}) \frac{k^t}{\sigma^t + 1} + (1 - \theta); \quad \gamma^t = \theta[\lambda^{t'} (k^{t'} + \delta^{t'}) + (1 - \lambda^{t'}) \gamma^t] \\ \delta^t &= \theta[\lambda^{t'} \gamma^{t'} + (1 - \lambda^{t'}) (k^t \frac{\sigma^t}{\sigma^t + 1} + \delta^t)]; \quad s^t = \theta[\lambda^{t'} s^{t'} + (1 - \lambda^{t'}) s^t] \end{aligned} \quad (13)$$

for $t \in \{A, B\}$. System (13) has eight equations and eight unknowns. Solving out, we obtain (7). **Uniqueness of linear equilibrium.** The proof is standard but lengthy, and is therefore relegated to the supplementary material. ■

Proof of lemma 4. We have

$$E[U(a_i - \alpha_i, a_i - \bar{a}_i) \mid t_i, x_i] = \theta E[(a_i - \bar{a}_i)^2 \mid t_i, x_i] + (1 - \theta) E[(a_i - \alpha_i)^2 \mid t_i, x_i]. \quad (14)$$

We now consider each element of (14) in turn. Substituting for a_i from (6), we have $a_i - \alpha_i = \theta(\bar{a}_i - \alpha_i)$. Hence, $E[(a_i - \alpha_i)^2 | t_i, x_i] = \theta^2 E[(\bar{a}_i - \alpha_i)^2 | t_i, x_i]$. Consider now the first term in (14). Substituting for a_i from (6), we have $a_i - \bar{a}_i = \bar{\bar{a}}_i - \bar{a}_i - (1 - \theta)(\bar{a}_i - \alpha_i)$. Hence,

$$E[(a_i - \bar{a}_i)^2 | t_i, x_i] = E[(\bar{\bar{a}}_i - \bar{a}_i)^2 | t_i, x_i] + (1 - \theta)^2 E[(\bar{a}_i - \alpha_i)^2 | t_i, x_i] - 2(1 - \theta) E[(\bar{\bar{a}}_i - \alpha_i)(\bar{a}_i - \bar{a}_i) | t_i, x_i]. \quad (15)$$

The last term in (15) can equivalently be expressed as $E(\bar{\bar{a}}_i^2 | t_i, x_i) - E(\bar{\bar{a}}_i \bar{a}_i | t_i, x_i) - E(\alpha_i \bar{\bar{a}}_i | t_i, x_i) + E(\alpha_i \bar{a}_i | t_i, x_i)$. By iterated expectations, $E(\bar{\bar{a}}_i \bar{a}_i | t_i, x_i) = E[E(\bar{\bar{a}}_i \bar{a}_i | t_i, \alpha_i, x_i) | t_i, x_i] = E[\bar{\bar{a}}_i E(\bar{a}_i | t_i, \alpha_i, x_i) | t_i, x_i] = E(\bar{\bar{a}}_i^2 | t_i, x_i)$ since $\bar{\bar{a}}_i = E(\bar{a}_i | t_i, \alpha_i, x_i)$. Similarly, $E(\alpha_i \bar{\bar{a}}_i | t_i, x_i) = E[\alpha_i E(\bar{a}_i | t_i, \alpha_i, x_i) | t_i, x_i] = E(\alpha_i \bar{a}_i | t_i, x_i)$. This proves that the last term in (15) is equal to zero. Putting these together we obtain (8). ■

In the proofs that follow, it will be convenient to use the following lemma.

Lemma 5. *Consider a manager i of group $t \in \{A, B\}$. For given (α_i, x_i) the manager's expectation over \bar{a}_i is equal to*

$$\bar{\bar{a}}_i = b_0 \alpha_i + \lambda^{t'} x_i + (1 - b_0 - \lambda^{t'}) \mu^t \quad (16)$$

where $b_0 \equiv (1 - \lambda^{t'}) (1 - \theta) / (1 - \theta + \sigma^t + \theta \lambda^{t'})$.

Proof. From the proof of proposition 3, $\bar{\bar{a}}_i = E(a_j | t, \alpha_i, x_i)$ for a generic $j \neq i$. From lemma 3,

$$E(a_j | t, \alpha_i, x_i) = \sum_{\tau=A,B} \lambda^\tau [k^\tau E(\alpha_j^\tau | t, \alpha_i, x_i) + \theta \lambda^{\tau'} E(x_j^\tau | t, \alpha_i, x_i) + (1 - k^\tau - \theta \lambda^{\tau'}) E(\mu^\tau | t, \alpha_i, x_i)] \quad (17)$$

where $\tau' \equiv \{A, B\} \setminus \tau$. Substituting for k^t and $k^{t'}$, for $E(\alpha_j^{t'} | t, \alpha_i, x_i) = E(\mu^{t'} | t, \alpha_i, x_i) = E(x_j^t | t, \alpha_i, x_i) = x_i$, and for $E(\alpha_j^t | t, \alpha_i, x_i) = (\sigma^t \mu^t + \alpha_i) / (\sigma^t + 1)$, we obtain (16). ■

Proof of proposition 1. Substituting for $\bar{\bar{a}}_i$ from lemma 5 above, $\bar{\bar{a}}_i - \alpha_i$ becomes $\alpha_i (b_0 - 1) + \lambda^{t'} x_i + (1 - b_0 - \lambda^{t'}) \mu^t = (e^t + \varepsilon_i) (b_0 - 1) + \lambda^{t'} (x_i - \mu^t)$ since $\alpha_i = \mu^t + e^t + \varepsilon_i$. It follows that $E[(\bar{\bar{a}}_i - \alpha_i)^2 | t, x_i] = (1 - b_0)^2 (1 + \sigma^t) + (\lambda^{t'})^2 (x_i - \mu^t)^2$ and $\frac{\partial E[(\bar{\bar{a}}_i - \alpha_i)^2 | t, x_i]}{\partial \lambda^{t'}} = 2[\lambda^{t'} (x_i - \mu^t)^2 + \frac{(\sigma^t + 1)^2 (1 - \theta)(\sigma^t + \lambda^{t'})}{(1 - \theta + \sigma^t + \theta \lambda^{t'})^3}] > 0$, which proves the result. ■

Proof of proposition 2. We have derived $E[(\bar{\bar{a}}_i - \alpha_i)^2 | t, x_i]$ above in the proof of proposition 1. Consider now $E[(\bar{\bar{a}}_i - \bar{a}_i)^2 | t, x_i]$. First, note that by iterated expectations, as in the proof of lemma 4, $E[(\bar{\bar{a}}_i - \bar{a}_i)^2 | t, x_i] = E(\bar{\bar{a}}_i^2 | t, x_i) - E(\bar{\bar{a}}_i^2 | t, x_i)$. From lemma 5 above,

$$E(\bar{\bar{a}}_i^2 | t, x_i) = E[(b_0 \alpha_i + \lambda^{t'} x_i + (1 - b_0 - \lambda^{t'}) \mu^t)^2 | t, x_i]. \quad (18)$$

Substituting for $\alpha_i = \mu^t + e^t + \varepsilon_i$ and solving, we obtain

$$E(\bar{a}_i^2 \mid t, x_i) = [\mu^t(1 - \lambda^{t'}) + x\lambda^{t'}]^2 + b_0^2(1 + \sigma^t). \quad (19)$$

Consider now $E(\bar{a}_i^2 \mid t, x_i)$. From the definition of \bar{a}_i , it is straightforward that $\bar{a}_i^2 = \frac{1}{n^2} \sum_{j=1}^n a_j^2 + \frac{2}{n^2} \sum_{j=1}^n \sum_{l \neq j} a_j a_l$. As there are $n(n-1)/2$ cross products, taking expectations, we obtain

$$E(\bar{a}_i^2 \mid t, x_i) = \frac{1}{n} [E(a_j^2 \mid t, x_i) - E(a_j a_l \mid t, x_i)] + E(a_j a_l \mid t, x_i) \quad (20)$$

for generic $j \neq i, l \neq j, i$ where $j, l \in \{1, 2, \dots, n\}$. Consider first $E(a_j^2 \mid t, x_i)$. From lemma 3, this is equal to

$$E(a_j^2 \mid t, x_i) = \sum_{\tau=A,B} \lambda^\tau E[(k^\tau \alpha_j^\tau + \theta \lambda^{\tau'} x_j^\tau + (1 - k^\tau - \theta \lambda^{\tau'}) \mu^\tau)^2 \mid t, x_i] \quad (21)$$

where $\tau' \equiv \{A, B\} \setminus \tau$. To evaluate the expectation in (21), which is conditional on $t_i = t$ and x_i , we first substitute for $\alpha_j^\tau = \mu^\tau + e^\tau + \varepsilon_j$ and then take expectations. We then further substitute for $E(\mu^{t'} \mid t, x_i) = x_i$, $E[(\mu^{t'})^2 \mid t, x_i] = x_i^2 + X$, $E[(x_j^{t'})^2 \mid t, x_i] = (\mu^t)^2 + X$, $E[(x_j^t)^2 \mid t, x_i] = E[(x_i - \xi_i + \xi_j)^2 \mid t, x_i] = x_i^2 + 2X$, for k^t and $k^{t'}$ and finally for $\lambda^t = 1 - \lambda^{t'}$. This allows to express $E(a_j^2 \mid t, x_i)$ as a function of $x_i, \mu^t, X, \sigma^t, \sigma^{t'}$ and $\lambda^{t'}$. Consider now $E(a_j a_l \mid t, x_i)$. This is equal to

$$(\lambda^{t'})^2 E(a_j^{t'} a_l^{t'} \mid t, x_i) + (1 - \lambda^{t'})^2 E(a_j^t a_l^t \mid t, x_i) + 2\lambda^{t'}(1 - \lambda^{t'}) E(a_j^t a_l^{t'} \mid t, x_i). \quad (22)$$

We first substitute for a_j, a_l (from lemma 3) and for $\alpha_j^t, \alpha_l^t, \alpha_j^{t'}, \alpha_l^{t'}, k^t, k^{t'}$, and take expectations. We then substitute for $E(x_j^t x_l^t \mid t, x_i) = E(x_j^t \mu^{t'} \mid t, x_i) = x_i^2 + X$, $E(x_j^{t'} \mu^{t'} \mid t, x_i) = E(x_j^t \mu^t \mid t, x_i) = E(\mu^t \mu^{t'} \mid t, x_i) = x_i \mu^t$, $E(x_j^{t'} x_l^{t'} \mid t, x_i) = E(x_j^{t'} \mu^t \mid t, x_i) = (\mu^t)^2$. Solving out, we obtain $E(a_j a_l \mid t, x_i)$ as a function of $x_i, \mu^t, X, \sigma^t, \sigma^{t'}$ and $\lambda^{t'}$. Overall, this allows to calculate $E[(\bar{a}_i - \bar{a}_i)^2 \mid t, x_i]$ and, thus, $E[U(\cdot) \mid t, x_i]$, as a function of these parameters. Let $T \equiv \lim_{\lambda^{t'} \rightarrow 0} \frac{\partial(E[U(\cdot) \mid t, x_i])}{\partial \lambda^{t'}}$. Straightforward computations then allow to derive

$$T = \frac{\theta}{2n} \{ (2\theta^2 - 2\theta + 1)X + (\theta - 1)^2 z_i^2 - \frac{(1 - \theta)^2}{(\sigma + 1 - \theta)^3} \phi \} \quad (23)$$

where $\phi \equiv (\sigma^{t'} + 1)\theta^3 - 3(\sigma^t + 1)(\sigma^{t'} + 1)\theta^2 + (\sigma^t + 1)^2(3\sigma^{t'} + \sigma^t + 4)\theta - (\sigma^t + 1)^3(\sigma^{t'} - \sigma^t + 2)$. Letting $\sigma^{t'} = \sigma^t + \Delta$, we obtain $\lim_{\theta \rightarrow 1} \phi = \sigma(3\sigma^t + (2 - \Delta)(\sigma^t)^2 + 1) > 0$ for Δ sufficiently small. This proves the result. ■

Proof of proposition 3. Following the method outlined in the proof of proposition 2, we can derive $E[U(\cdot) | t, x_i]$ as a function of $z_i, X, \sigma^t, \sigma^{t'}$ and $\lambda^{t'}$. Let $Q \equiv \frac{d(\lim_{n \rightarrow \infty} E[U(\cdot) | t, x_i])}{d\lambda^{t'}}$. Straightforward computation then show that

$$Q = \lambda^{t'} \theta \{X + \theta(1 - \theta) z_i^2 + (1 - \theta)^2 \frac{\theta \lambda^{t'} (\sigma^{t'} - \sigma^t) + (\sigma^{t'} + 1)(2 - \theta + 2\sigma^t)}{(\sigma^{t'} + 1 - \theta \lambda^{t'})^3 (\sigma^t + 1 - \theta + \theta \lambda^{t'})^3} (\theta^2 \Lambda_2 - \theta \Lambda_1 + \Lambda_0)\} \quad (24)$$

where $\Lambda_2 \equiv [(\sigma^{t'})^2 + (\sigma^t + \sigma^{t'}) (\sigma^t + 3) + 3](\lambda^{t'})^2 - (\sigma^{t'} + 1)(2\sigma^{t'} + \sigma^t + 3)\lambda^{t'} + (\sigma^{t'} + 1)^2$, $\Lambda_1 \equiv (\sigma^{t'} + 1)(\sigma^t + 1)[\sigma^{t'}(1 - \lambda^{t'}) + \sigma^t \lambda^{t'} + 1]$ and $\Lambda_0 \equiv (\sigma^{t'} + 1)^2 (\sigma^t + 1)^2$. Note that $\Lambda_0 - \Lambda_1 > 0$ and, hence, $\Lambda_0 - \theta \Lambda_1 > 0$.²⁴ Consider now Λ_2 . This is a convex function of $\lambda^{t'}$. Evaluated at its minimum value w.r.t. $\lambda^{t'}$, $\Lambda_2 > 0$.²⁵ This proves that $\theta^2 \Lambda_2 - \theta \Lambda_1 + \Lambda_0 > 0$ and, hence, that $Q \geq 0$ (with strict inequality for $\lambda^{t'} > 0$). ■

Proof of proposition 4. From the proof of proposition 2 we know that $E[(\bar{a}_i - \bar{a}_i)^2 | t, x_i]$ can be written as

$$\frac{1}{n} [E(a_{j2}^2 | t, x_i) - E(a_{j1} a_l | t, x_i)] + E(a_{j1} a_l | t, x_i) - E(\bar{a}_i^2 | t, x_i). \quad (25)$$

Consider first $E(a_{j1} a_l | t_i, x_i) - E(\bar{a}_i^2 | t_i, x_i)$. Using (22) and (19), it is straightforward to see that this is equal to $(\lambda^{t'})^2 X + (\lambda^{t'} k^{t'})^2 + \sigma^t [(1 - \lambda^{t'}) k^t]^2$. Remark that each element of this expression is convex in $\lambda^{t'}$, and, thus, their sum is convex. To prove that $E[(\bar{a}_i - \bar{a}_i)^2 | t, x_i]$ is convex in $\lambda^{t'}$, it is then sufficient to show that this holds when $n = 1$. This follows since, when $n = 1$, the weight assigned to the expression in square brackets in (25) is largest. When $n = 1$, $E[(\bar{a}_i - \bar{a}_i)^2 | t, x_i]$ can be written as

$$\lambda^{t'} (\sigma^{t'} + 1) (k^{t'})^2 + \Omega + \frac{1 - \lambda^{t'}}{\sigma^t + 1} \left(2\sigma^t + \lambda^{t'} + (\sigma^t)^2 \right) (k^t)^2 \quad (26)$$

where $\Omega \equiv \lambda^{t'} [1 - 2(1 - \lambda^{t'}) \theta (1 - \theta)] X + \lambda^{t'} (1 - \lambda^{t'}) (1 - \theta)^2 z_i^2$. The first component is convex in $\lambda^{t'}$. The second derivative of the second component of (26) w.r.t. $\lambda^{t'}$ is $2(1 - \theta) [2\theta X - z_i^2 (1 - \theta)] > 0$ for z_i sufficiently small. Finally, the second derivative of the third component of (26) w.r.t. $\lambda^{t'}$ is $2(1 - \theta)^2 \frac{(\sigma^t + 1)^3}{(\sigma^t - \theta + \theta \lambda^{t'} + 1)^4} \Pi$, where $\Pi \equiv (1 - \lambda^{t'}) (\sigma^t + 1) \theta^2 + (2(\sigma^t)^2 + 4\sigma^t + 2\lambda) \theta - (\sigma^t + 1)$. Note that $\lim_{\theta \rightarrow 1} \Pi = \sigma^t (4 - \lambda^{t'}) + \lambda^{t'} + 2(\sigma^t)^2 - \sigma^t \lambda^{t'} > 0$. This proves the result. ■

²⁴More precisely, $\Lambda_0 - \Lambda_1 = (\sigma^t + 1)(\sigma^{t'} + 1)[\sigma^t(1 - \lambda^{t'}) + \sigma^{t'}(\sigma^t + \lambda^{t'})] > 0$.

²⁵More precisely, $\Lambda_2 = \frac{3}{4} \frac{(\sigma^{t'} + 1)^2 (\sigma^t + 1)^2}{3\sigma^{t'} + 3\sigma^t + (\sigma^{t'})^2 + (\sigma^t)^2 + \sigma^{t'} \sigma^t + 3} > 0$.

Proof of proposition 5. In order to compute $E[Var(a_j | e^A, e^B) | t_i, x_i]$, we first derive $Var(a_j | e^A, e^B)$ and then take expectations conditional on t_i and x_i . Note that $Var(a_j | e^A, e^B) = E[(a_j - E(a_j | e^A, e^B))^2 | e^A, e^B] = E(a_j^2 | e^A, e^B) - [E(a_j | e^A, e^B)]^2$. Consider first $E(a_j^2 | e^A, e^B)$. Substituting for a_j from lemma 3, substituting for $\alpha_j^\tau = \mu^\tau + e^\tau + \varepsilon_j$, $x_j^\tau = \mu^{\tau'} + \xi_j$ and taking expectations, we obtain

$$E(a_j^2 | e^A, e^B) = \sum_{\tau=A, B, \tau' \equiv \{A, B\} \setminus \tau} \lambda^\tau [(\theta \lambda^{\tau'})^2 X + (k^\tau)^2 ((e^\tau)^2 + \sigma^\tau)] + \lambda^B [\theta \lambda^A \mu^A + (1 - \theta \lambda^A) \mu^B]^2 + \lambda^A [\theta \lambda^B \mu^B + (1 - \theta \lambda^B) \mu^A]^2. \quad (27)$$

Consider now $[E(a_j | e^A, e^B)]^2$. This is equal to

$$[E(a_j | e^A, e^B)]^2 = [\lambda^B E(a_j^B | e^A, e^B) + \lambda^A E(a_j^A | e^A, e^B)]^2. \quad (28)$$

Substituting for a_j from lemma 3, $[E(a_j | e^A, e^B)]^2$ can be written as

$$\{\theta \mu^A \lambda^B \lambda^A + \theta \mu^B \lambda^A \lambda^B + \sum_{\tau=A, B, \tau' \equiv \{A, B\} \setminus \tau} \mu^\tau \theta \lambda^{\tau'} (1 - \gamma^\tau) + \lambda^\tau k^\tau e^\tau\}^2. \quad (29)$$

$Var(a_j | e^A, e^B)$ is obtained by subtracting (29) from (27). After rearranging this gives

$$\lambda^A \lambda^B (\theta - 1)^2 (\mu^A - \mu^B)^2 + \theta^2 \lambda^A \lambda^B X + \sum_{\tau=A, B} \frac{\lambda^\tau (1 - \theta)^2 (\sigma^\tau + 1)^2}{(\sigma^\tau - \theta \lambda^\tau + 1)^2} [(e^\tau)^2 + \sigma^\tau]. \quad (30)$$

It remains to compute the expectation of (30) conditional on $t_i = t$ and x_i . After substituting for $E[\mu^{\tau'} | t, x_i] = x_i$, $E[(\mu^{\tau'})^2 | t, x_i] = x_i^2 + X$, $E[(e^\tau)^2 | t, x_i] = 1$ for $\tau = A, B$, and for $\lambda^t = 1 - \lambda^{t'}$ we obtain $E[Var(a_j | e^A, e^B) | t, x_i]$ as a function of x_i , μ^t , X , σ^t , $\sigma^{t'}$ and $\lambda^{t'}$. Let $P \equiv \lim_{\lambda^{t'} \rightarrow 0} \frac{\partial E[Var(a_j | e^A, e^B) | t, x_i]}{\partial \lambda^{t'}}$. Straightforward calculations show that

$$P = (\theta - 1)^2 z_i^2 + (-2\theta + 2\theta^2 + 1) X - \frac{(\theta - 1)^2}{(1 - \theta + \sigma^t)^3} \Psi \quad (31)$$

where $\Psi \equiv \theta^3 (\sigma^{t'} + 1) - 3\theta^2 (\sigma^{t'} + 1)(\sigma^t + 1) + \theta (\sigma^t + 1)^2 (3\sigma^{t'} + \sigma^t + 4) - (\sigma^t + 1)^3 (\sigma^{t'} - \sigma^t)$. Letting $\sigma^{t'} = \sigma^t + \Delta$, we obtain $\lim_{\theta \rightarrow 1} \Psi = 7\sigma^t + 9(\sigma^t)^2 + (4 - \Delta)(\sigma^t)^3 + 2 > 0$ for Δ sufficiently small. This proves the result. ■